



On the Design of Hierarchies: Coordination Versus Specialization

Citation

Hart, Oliver D., and John Moore. 2005. On the design of hierarchies: Coordination versus specialization. *Journal of Political Economy* 113(4): 675-702.

Published Version

doi:10.1086/431794

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3448676>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

On the Design of Hierarchies: Coordination versus Specialization

Oliver Hart

Harvard University and London School of Economics

John Moore

Edinburgh University and London School of Economics

We consider an economy that has to decide how assets are to be used. Agents have ideas, but these ideas conflict. We suppose that decision-making authority is determined by hierarchy: each asset has a chain of command, and the most senior person with an idea exercises authority. We analyze the optimal hierarchical structure given that some agents coordinate and other specialize. Among other things, our theory explains why coordinators should typically be senior to specialists and why pyramidal hierarchies may be optimal. Our theory also throws light on the optimal degree of decentralization inside a firm and on firm boundaries.

We would like to acknowledge the very helpful suggestions of Philippe Aghion, Carliss Baldwin, Antoine Faure-Grimaud, Bob Gibbons, Louis Kaplow, Julio Rotemberg, Andrea Prat, Ilya Segal, editors Steve Levitt and Canice Prendergast, and an anonymous referee and the research assistance of Ulrike Malmendier. We have also benefited from comments from Abhijit Banerjee, Patrick Bolton, Rebecca Henderson, Arthur Hosios, Bentley MacLeod, Fausto Panunzi, David Parkes, Rohan Pitchford, Andrew Postlewaite, Michael Schwarz, Andrei Shleifer, Jean Tirole, Birger Wernerfelt, and Chenggang Xu and from the reactions of seminar audiences at the Western Economic Association and European Association for Research in Industrial Economics meetings. We acknowledge financial support from the U.S. National Science Foundation through the National Bureau of Economic Research and from the U.K. Economic and Social Research Council.

[*Journal of Political Economy*, 2005, vol. 113, no. 4]
© 2005 by The University of Chicago. All rights reserved. 0022-3808/2005/11304-0001\$10.00

I. Introduction

Suppose that a decision must be made in an organization. The organization might represent a private firm, a public firm, or even society. Assume that individual i favors one decision and individual j favors another. Who gets his or her way? In many cases, the answer is that this depends on who is senior in the organization. That is, there is a hierarchy of individuals with respect to any decision. In the event of a conflict, the decision will be made by the most senior person in the hierarchy who cares about—or has a view about—this particular decision.

In this paper we develop a model of hierarchies. We consider an economy that has to decide how assets are to be used. Agents may have ideas about asset usage, but these ideas conflict; that is, only one person's idea can be implemented with respect to any one asset. We suppose that each asset has a chain of command, and the most senior person in the chain who has an idea—who needs access to the asset—implements her idea. The order in which ideas are proposed does not matter (“top down” and “bottom up” both lead to the same outcome) given that a senior person is willing to defer to her juniors if she does not have an idea herself.

We use our model to analyze the optimal hierarchical structure given that different agents have different tasks; in particular, some agents are engaged in coordination and others in specialization. Among other things, our theory explains why coordinators should typically be senior to specialists and why pyramidal hierarchies may be optimal. Our theory also throws light on the optimal degree of decentralization inside a firm and on the boundaries of the firm.

Before proceeding, we should acknowledge that the assumption that decisions are made hierarchically is hardly uncontroversial. For example, the Coase theorem suggests that the parties should bargain about which decision is made and divide up the surplus in some appropriate manner. Alternatively, the right to have access to an asset could be auctioned off or decisions could be determined by voting. In this paper we put the Coase theorem, auction mechanisms, and voting to one side and assume that decisions are made hierarchically. However, we offer some justifications for this assumption. We describe some conditions under which a hierarchy may be the only feasible organizational structure. Also, we show that under some conditions a hierarchy achieves first-best, in which case other organizational forms do not have to be considered.

To give a flavor of the analysis, consider the case of two assets, a_1 and a_2 . One agent specializes in thinking of ideas about how to use a_1 ; another agent specializes on a_2 . A third agent coordinates the two assets: she thinks of ideas about how the assets may be used in tandem. The

question is, if the coordinator's ideas conflict with those of the specialists, whose idea should take priority?

The most natural hierarchy has the coordinator senior on both assets. In this case, if the coordinator has an idea, she will implement it, whether or not the specialists have ideas, since she has authority. In contrast, each specialist will implement his idea only if (a) he has one and (b) the coordinator does not.

Interestingly, this structure is not always best. It may be dominated by another hierarchy in which the coordinator is *junior* to the specialist on each asset. In this case, each specialist will implement an idea whenever he has one, whereas the coordinator will implement her idea only if (a) she has one and (b) neither specialist does. (If either specialist implements his idea, this preempts the use of one of the assets, which means that the coordinator cannot implement her idea, given that she needs to use the assets in tandem.) Under what circumstances can we rule out these less natural hierarchies?

One of our main propositions makes use of a fairly plausible monotonicity condition, that an agent's probability of having an idea decreases with her span of control. Assuming that this condition holds, we prove that agents with a broad remit, that is, whose tasks cover a large subset of assets, should appear higher in the chain of command than those with a narrow remit. We interpret agents with a broad remit to be strategic planners, who contemplate future directions for the organization. So our result says that big thinkers should be senior to small thinkers: coordinators should be senior to specialists. The intuition for this result is the following. A coordinator, who has an idea comparatively rarely, appears at all in an organization only because she has the potential to generate large value. But to do so, she must have control over all the assets that she needs. In other words, she must be senior: she must not be blocked on some key asset by a specialist.

We also establish that "crisscross" hierarchies are never optimal; that is, if individual i appears above individual j on one asset, j will not appear above i on another asset. The reason for this is that if i is senior to j on one asset and j is senior to i on another, then i and j will block each other if both of them have an idea.

Under an additional assumption that synergies are nested, we can go further and show that the optimal hierarchy is a pyramid, in the sense that if an agent has a boss on one asset, he has the same boss on every asset he works on.

Our model throws light on the trade-off between centralization and decentralization, as well as on the boundaries of the firm. The terms "centralization" and "decentralization" are often employed in the literature, but their meaning is sometimes not very clear. We define an organization to be centralized if it is likely that a decision will be made

by someone at headquarters (a coordinator) rather than by someone with local expertise (a specialist). We define an organization to be decentralized if a decision is likely to be made by a specialist rather than by a coordinator. One advantage of our approach is that centralization, decentralization, and independent firms are points on a continuum and the trade-off between them can be analyzed in a single model. We show that, if the gains to coordination are large, it is optimal for the organization to be centralized; if the gains to coordination are moderate, it is optimal for the organization to be decentralized (in a class of cases); and finally, if the gains to coordination are small, then it is optimal for the organization to split up into several independent firms (independent firms correspond to the case in which two assets have different people at the top of their chains of command, i.e., different owners).

On a very different tack, we think that our analysis may relate to the design of computer operating systems. When multiple central processor units (CPUs) operate in parallel and programs need the use of more than one processor at a time, instantaneous ranking decisions must be made to adjudicate the competing demands of different programs. Should a program that simultaneously needs several CPUs have priority over a group of programs that need only one CPU each? This is akin to our question of whether a coordinator should be senior to a group of specialists: the use of a CPU is analogous to access to an asset. The analogy is not exact, not least because our model of hierarchies is static (if one individual has access to an asset, then all other individuals are precluded from that asset), whereas a computer operating system is dynamic (running one program on a CPU merely delays another program from being run on that CPU).¹ There is one sense in which the design of a hierarchy for a computer operating system may be a simpler problem than ours, though: CPUs are generic (a program can typically run on any subset of the CPUs), whereas in our model an asset is specific to the agents whose tasks use it. For more on the design of hierarchies in parallel-processing computer systems, see Tanenbaum (2001, 521–26).

There is a vast literature on many of the issues we consider, and this is not the place to provide a review. Economists have studied hierarchical structure from the following points of view: productivity enhancement and span of control (see, e.g., Williamson 1967), optimal monitoring (see, e.g., Calvo and Wellisz 1979), task assignment with or without coordination (see, e.g., Rosen 1982; Harris and Raviv 2002), information processing and team theory (see, e.g., Keren and Levhari 1979; Radner 1992; Bolton and Dewatripont 1994; Van Zandt 1999), management by

¹ For a discussion of mechanisms such as hierarchy for allocating a scarce resource in a dynamic context, see Lando (2004).

exception (see, e.g., Garicano 2000), and optimal statistical decision making (see, e.g., Sah and Stiglitz 1986). By and large, however, with the exception of papers by Aghion and Tirole (1997), Baker, Gibbons, and Murphy (1999), and Rajan and Zingales (2001), discussed further below, the existing literature does not analyze hierarchy in terms of authority; that is, it is not the case that if individual i is above individual j in the hierarchy, i necessarily has authority over j . Rather, in much of the literature, if i is above j in the hierarchy, then j provides information to i . Also, the literature does not distinguish between what happens inside a firm and what happens between firms.

The papers by Aghion and Tirole (1997), Baker et al. (1999), and Rajan and Zingales (2001) *do* study hierarchical structure in terms of authority (or power). Aghion and Tirole study a two-person hierarchy in which one person is the boss and the other the subordinate. Each can have an idea. The boss's idea is implemented if she has one, whereas the subordinate's idea is implemented if he has one but the boss does not. Our paper is close in spirit to Aghion and Tirole's; the main differences are that they emphasize the ex ante incentive to have ideas rather than ex post efficiency, and they do not consider multiperson hierarchies. Baker et al. consider a repeated version of Aghion and Tirole and study how the boss may develop a reputation for implementing the subordinate's ideas even when, from a short-term perspective, she might prefer to implement her own. Finally, Rajan and Zingales consider a model in which the hierarchical structure determines who interacts with whom. The idea is that under a large, steep hierarchy the boss is insulated from many of the firm's workers and knows little about them, which gives the boss's close subordinates considerable bargaining power. The opposite is true for flat hierarchies. Rajan and Zingales draw various conclusions about optimal organizational form in physical capital versus human capital industries.

The paper is organized as follows. In Section II we present the model. Sections III and IV show how the model works in the case of two agents and three agents, respectively. Section V contains general results for the m -asset, n -agent case. In Section VI we analyze the trade-off between centralization and decentralization. Section VII provides some interpretation of the model and discusses certain foundational issues. Finally, Section VIII presents conclusions.

II. The Model

We consider a two-date economy consisting of m assets, a_1, \dots, a_m , and n risk-neutral agents 1, ..., n . Each asset represents a decision that has to be made ex post, at date 1. These decisions are noncontractible both ex ante and ex post. However, a hierarchy of authority over decisions,

an organizational form, can be contractually specified *ex ante*, at date 0.

Agents are assigned tasks at date 0. A task consists of thinking about what to do with a particular subset of the m assets, that is, what decisions to make with respect to these assets at date 1. For each subset $A \subset \{a_1, \dots, a_m\}$, there is a corresponding task $t(A)$.² More than one person can be assigned the same task.

A task $t(A)$ may or may not reach fruition at date 1 in that with some probability $0 < p(A) < 1$ the agent carrying out the task can generate value $v(A) \geq 0$ (as long as he or she has access to all the assets in A), but with probability $1 - p(A)$, no value can be generated.³ We refer to the event in which value can be generated as “the agent has an idea.” We make the following assumption.

ASSUMPTION 1. All ideas are stochastically independent events.

If two or more people are assigned the same task $t(A)$, then each, independently, has a probability $p(A)$ of having an idea.

We put no restrictions on the valuation function $v(\cdot)$. It may not be superadditive or even nondecreasing in A . We have in mind a situation in which thinking about how to use two assets is a very different activity from thinking about how to use one of them. The first activity involves coordination whereas the second does not. If coordination possibilities are limited, then the value of having an idea about how to coordinate two assets may be very low, or even zero. If a coordinator is to succeed in generating any value at date 1, we assume that he must have access to—have decision-making power over—all the assets in his remit.

ASSUMPTION 2. An agent carrying out task $t(A)$ who has an idea generates value $v(A)$ if and only if he has access to *all* the assets in A .

This is an important restriction. Although there is no loss of generality in defining a task by the set of assets needed to generate value, it is restrictive to limit each agent to one task.⁴ We shall return to this point in Section VIII.

If two people have an idea as to how to use an asset, they cannot both have their way.

ASSUMPTION 3. Only one person can have access to, or use of, an asset at a time.

This is the source of tension in the model: the asset is a scarce re-

² The model is easily generalized to have more than one possible task for each subset of assets A , although this would complicate the notation.

³ More generally, we might assume that a task $t(A)$ always generates some positive baseline value and that, with probability $p(A)$, an additional value $v(A)$ can be generated. Note that there must be some subsets A for which $v(A)$ is strictly positive; otherwise the problem is uninteresting.

⁴ The model could be extended to allow an agent to undertake more than one task—perhaps with a concomitant reduction in the probability of his having an idea on each—but a number of our results would have to be modified.

source. At the heart of our analysis lies the question, how are conflicts over the use of an asset to be resolved?

We assume that at date 0 an organizational form is put in place that determines a chain of command for how each asset is to be used at date 1.

ASSUMPTION 4. Access to assets is determined by a hierarchical structure. That is, each asset a_k has a chain of command, a list L_k , that ranks agents by seniority over that asset.

In principle, there are many nonhierarchical structures that we might also consider, such as voting rules or the use of mediators to resolve conflicts. But for the moment let us suppose that all such arrangements are infeasible. This matter is taken up further in Section VII.

People without ideas at date 1 are passive and drop out of the picture.

ASSUMPTION 5. On each asset a_k , senior agents who do not have an idea do not exercise authority; they defer to the highest-ranking person on list L_k who does have an idea.

However, someone with an idea is *always* active, even if, because of assumption 2, he is unable to generate any value.

ASSUMPTION 6. The most senior person with an idea exercises authority whether or not he has access to the other assets he needs to generate value. There is no ex post renegotiation of authority at date 1.

Assumption 6, in conjunction with assumptions 2 and 4, is perhaps the strongest of our assumptions.⁵

In contrast to assumption 6, we make the following assumption.

ASSUMPTION 7. The assignment of tasks to agents and the hierarchical structure are chosen at date 0 to maximize expected value.

We have in mind that at date 0 the parties are able to negotiate, with the surplus being divided up using lump-sum transfers.

A number of interpretations of the model are offered in Section VII. We shall maintain all seven assumptions, 1–7, throughout. For most of the paper we shall make an eighth assumption.

ASSUMPTION 8. All agents have an equal ability to undertake any task.

Assumption 8 implies that in the date 0 assignment of tasks, it does not matter who does what; it only matters which tasks are undertaken. As we shall see, this assumption is useful because some—but by no means all—of our general results depend on the ability to change the composition of tasks that are undertaken. The alternative to assumption 8 would be to allow for skill specificity across agents, and, in the extreme, task assignment might even be fixed exogenously. In Sections III and

⁵ An interesting open question is how far our results would be affected by instead assuming that people exercise authority only if they can generate value; we conjecture that, in substance, our results would remain intact.

IV, we shall find it helpful to consider the case of fixed tasks as a preliminary exercise before dealing with the case of endogenous task assignment.

Before we write down the formula for date 0 expected total value, we can simplify matters a little. Clearly, there is no point in having anyone entered more than once on some list L_k , because only the first, most senior, entry would be pertinent. Also, if an agent's task consists of looking after assets in the subset A , then it makes no sense to put the agent on the list L_k pertaining to an asset $a_k \notin A$, since the agent will never have an idea about a_k . Similarly, it makes no sense *not* to put him on the list L_k pertaining to each asset $a_k \in A$, since by assumption 2 he generates no value unless he has access to all the assets in A .

Putting these observations together, we can conclude the following. Once the lists for all the assets have been determined, the task each person is doing can be deduced by seeing on which list he appears. Specifically, for agent i define

$$A_i = \{\text{assets } a_k | i \text{ appears on list } L_k\}.$$

Given that A_i is the maximal set of assets over which i can exercise authority, he must be engaged in task $t(A_i)$.

For each agent i , define

$$S_i = \{\text{agents } j | \text{for some asset } a_k, i \text{ and } j \text{ both appear on list } L_k \\ \text{and } j \text{ appears above } i\}.$$

The term S_i is the set of agents who are senior to i on some asset. Now we know from assumptions 2 and 6 that i generates value $v(A_i)$ if and only if he has an idea and nobody senior to him on any of the assets he looks after has an idea. We can therefore write the formula for the date 0 expected total value:

$$V = \sum_{i=1}^n p(A_i) \left\{ \prod_{j \in S_i} [1 - p(A_j)] \right\} v(A_i). \quad (1)$$

By assumption 7, an optimal organization form, the m lists, will be chosen to maximize (1).

III. Two Agents

Let us start with a simple case in which there are just two agents, 1 and 2. As a preliminary exercise, let us ignore the choice of task assignment, and suppose that they have fixed tasks $t(A_1)$ and $t(A_2)$, respectively. Given that their tasks are fixed, the only question is, which agent should be senior on each of the assets in $A = A_1 \cap A_2$, the assets they have in

common? Assume that A is nonempty (if A is empty, authority is not an issue). Note that seniority matters only in the event that both agents have an idea, because, by assumption 5, if only one has an idea, the other will defer irrespective of seniority. Also note that if both agents have an idea, they cannot both generate surplus because, by assumption 2, they would both require access to the assets in A at the same time, which is ruled out by assumption 3.

We make two claims about the optimal seniority structure.

CLAIM 1. It is never optimal to have a crisscross hierarchy in which 1 is senior to 2 on some asset in A and 2 is senior to 1 on another.

With a crisscross hierarchy, if both agents have an idea, then, by assumption 6, both will exercise authority on the assets on which they are senior and, by assumption 2, no value will be generated because they will block each other. As we shall prove later, claim 1 is a general result, for any number of agents.⁶

Our second claim is specific to the two-agent case.

CLAIM 2. With two agents, it is optimal for the agent with a higher value to be senior.

That is, if $v(A_1) > v(A_2)$, then 1 should be senior to 2 on all assets in A . Because seniority matters only if both agents have an idea, we can condition on this event. And given that more value is generated by 1, he should be senior.

From claim 2, it would appear that the probabilities $p(A_1)$ and $p(A_2)$ are irrelevant to the choice of hierarchy. But this is no longer true once we endogenize the agents' tasks. If we make assumption 8, then claim 2 can be reexpressed in terms of probabilities alone, without reference to values.

CLAIM 3. With endogenous task assignment, it is never optimal for an agent with a lower probability of having an idea to be junior.

This seemingly paradoxical claim is easy to demonstrate in the case of two agents. Suppose, contrary to claim 3, that 2 is junior to 1 in an optimal hierarchy but has a lower probability of having an idea: $p(A_2) < p(A_1)$. Then, since it cannot be the case that $v(A_2) > v(A_1)$ (otherwise, by claim 2, 2 should be senior to 1), 2's task is strictly dominated by 1's. Not only is she less likely to have an idea, but even when she has an idea she generates no more value. In which case, we can invoke assumptions 7 and 8 and conclude that 2 should be reassigned 1's task. That is, expected total value would increase by having both agents do task $t(A_1)$ —contrary to our hypothesis that the initial hierarchy is optimal.

Most surprisingly (at least to us), it turns out that this curious result—

⁶ If we had allowed agents to have more than one task (recall the discussion after assumption 2), then claim 1 would not hold in general (see Sec. VIII).

that in an optimal hierarchy agents' seniorities are inversely related to their probability of having an idea—is quite general and provides the key lemma from which all other results follow as corollaries.

Before leaving the two-person hierarchy, we should mention a useful analogy from tennis (suggested to us by Andrea Prat). A player usually has two kinds of serves: fast and slow. A fast serve has a lower chance of success but a higher impact if successful. On each service point, a player typically chooses to serve fast on his first serve and, if that fails, slow on his second. Sometimes he may choose to serve fast twice, or slow twice, but never slow then fast. This squares with claim 3: the fast serve (with the lower chance of success) is never “junior” to the slow serve.⁷

It is straightforward to confirm that, just as with the choice of serves in tennis, the design of an optimal two-person hierarchy typically has an “interior” solution; that is, the agents are assigned different tasks, even though at first glance the problem appears inherently linear. The point is that, for any task $t(A)$, the *option* value, $v(A)$, assumes importance in the design of an optimal hierarchy.⁸ By contrast, in a one-person problem only the *expected* value $p(A)v(A)$ would matter.

The case of two agents is too simple to allow us to get at interesting questions to do with coordination and specialization. We need to turn to three agents.

IV. Three Agents

Suppose now that there are three agents, 1, 2, and 3, and, for simplicity, just two symmetric assets, a_1 and a_2 . This means that there are only two kinds of tasks: an agent can be either a specialist, working on one of the assets, or a coordinator, working on both. Let a specialist have an idea with probability p_s , which generates value v_s ; and let a coordinator have an idea with probability p_c , which generates value v_c .⁹

Again, it is instructive, as a preliminary exercise, to fix the agents' tasks. Suppose that agents 1 and 2 must be specialists, working on assets a_1 and a_2 , respectively, and agent 3 must be a coordinator. Given this exogenous assignment of tasks, there are really only two possible hierarchies to consider: either the coordinator is senior to the specialist

⁷ If there were some benefit to winning the point on the first serve (e.g., if there were discounting), then this would not necessarily be true. The analogy with tennis is not exact because in our model the “first serve” (the senior agent) preempts, rather than precedes, the “second serve” (the junior agent).

⁸ This point is related to the observation of Roberts and Weitzman (1981) about the role of option value rather than expected return in determining the optimal sequence of exploration projects.

⁹ That is, $v(\{a_1\}) = v(\{a_2\}) = v_s$, $p(\{a_1\}) = p(\{a_2\}) = p_s$, and $v(\{a_1, a_2\}) = v_c$, $p(\{a_1, a_2\}) = p_c$.

TABLE 1

HIERARCHY <i>css</i> (Coordinator Senior)		HIERARCHY <i>ssc</i> (Coordinator Junior)	
Asset a_1	Asset a_2	Asset a_1	Asset a_2
3	3	1	2
1	2	3	3

on each asset or she is junior. Call these hierarchies *css* and *ssc*, respectively (see table 1). The third possibility, in which the coordinator is senior to agent 1 on asset 1 but junior to agent 2 on asset 2, or vice versa, is never optimal in this symmetric example, so we do not consider it.

The date 0 expected total values from the two hierarchies are, respectively,

$$V(css) = p_c v_c + (1 - p_c) 2p_s v_s \quad (2)$$

and

$$V(ssc) = 2p_s v_s + (1 - p_s)^2 p_c v_c \quad (3)$$

The first term in the expression for $V(css)$ comes from the fact that if the coordinator 3 is senior, then she generates expected value $p_c v_c$. It is only when, with probability $1 - p_c$, she does not have an idea that specialists 1 and 2 can each generate expected value $p_s v_s$; hence the second term in (2). The first term in the expression for $V(ssc)$, by contrast, comes from the fact that if 1 and 2 are senior, they can together generate $2p_s v_s$. If either of them has an idea, then, by assumption 2, they block 3 from access to a needed asset. It is only when, with probability $(1 - p_s)^2$, neither of them has an idea that 3 can generate expected value $p_c v_c$; hence the second term in (3).

From (2) and (3), we have

$$V(css) \geq V(ssc) \quad \text{as} \quad (2 - p_s) v_c \geq 2v_s. \quad (4)$$

At one extreme, if $v_c > 2v_s$, it is better to make the coordinator senior, irrespective of p_s , because the value of her idea is higher than the combined value of the specialists' ideas. At the other extreme, if $v_s > v_c$, it is better to make the specialists senior, irrespective of p_s , because the value of just one of their ideas is higher than that of the coordinator's idea. In the middle ground, where $2v_s > v_c > v_s$, there is a critical value of p_s : below this value, it is better to make the coordinator senior; above, it is better to make the specialists senior. The intuition is that seniority matters only when the coordinator and at least one of the specialists have an idea. Given $2v_s > v_c > v_s$, it is better for the coordinator's idea

TABLE 2

HIERARCHY <i>sss</i> (Three Specialists)		HIERARCHY <i>ccc</i> (Three Coordinators)	
Asset a_1	Asset a_2	Asset a_1	Asset a_2
1	2	1	1
3		2	2
		3	3

to be implemented if exactly one specialist has an idea and for the coordinator's idea not to be implemented if both specialists have ideas. The former event is relatively likely if p_s is low and relatively unlikely if p_s is high.

One lesson to be learned from this is that with three or more agents the choice of seniority, even with fixed tasks, depends on the probabilities p as well as on the values v —in contrast to the two-agent case with fixed tasks, where only the relative values mattered.¹⁰ That is, for three or more agents, there are no general results equivalent to claim 2.

So far in this section all we have done is to compare the two hierarchies in table 1, keeping tasks fixed. It is time to allow for an endogenous assignment of tasks. For the remainder of the paper, we shall assume that assumption 8 holds.

There are several possible task assignments and hierarchies to consider in addition to those in table 1, but since Section V deals with the general case, let us for now look at the two extreme hierarchies: hierarchy *sss*, comprising three specialists (with two agents specializing on, say, asset a_1), and hierarchy *ccc*, comprising three coordinators (see table 2).¹¹

The date 0 expected total values from these hierarchies are, respectively,

$$V(sss) = [1 - (1 - p_s)^2]v_s + p_s v_s \quad (5)$$

and

$$V(ccc) = [1 - (1 - p_c)^3]v_c \quad (6)$$

The first term in the expression for $V(sss)$ comes from the fact that v_s is realized on asset a_1 if either 1 or 3 has an idea. The expression for $V(ccc)$ comes from the fact that v_c is realized unless no one has an idea.

¹⁰ In a comparison of hierarchies *css* and *ssc*, the value of p_c is irrelevant because seniority matters only in the event that the coordinator has an idea. But this pertains to fixed tasks. As we shall see, when tasks are endogenous, both p_s and p_c are important in determining the optimal hierarchy.

¹¹ The cases we have not considered involve two coordinators and one specialist. But our conclusions hold even if these possibilities are allowed for.

The following results can be established by simple algebra.

RESULT 1. Suppose $p_c < p_s$. Then $V(ssc) < \max [V(css), V(sss)]$.

RESULT 2. Suppose $p_s < p_c$. Then $V(css) < \max [V(ssc), V(ccc)]$.¹²

Although we have not considered all possible hierarchies, results 1 and 2 suggest that in an optimal hierarchy an agent with a lower probability of having an idea is never junior, just as we found in claim 3 for the case of two agents. For example, in result 1, where the coordinator has a lower probability of having an idea, it is not optimal to make her junior to two specialists: hierarchy *ssc* is dominated by at least one other hierarchy. The rough intuition is that either v_c is high relative to v_s or it is low. If high, then the coordinator's idea should be implemented if she has one; that is, hierarchy *css* dominates hierarchy *ssc*. If low, then, given $p_c < p_s$, the coordinator is inferior in every way to a specialist—she has a smaller probability of having an idea and her idea has smaller value—and so she should be reassigned as a specialist; that is, hierarchy *sss* dominates hierarchy *ssc*.¹³ A broadly similar logic applies in result 2: either v_s is high relative to v_c and hierarchy *ssc* dominates hierarchy *css*, or v_s is low relative to v_c and, given $p_s < p_c$, the two specialists should be reassigned as coordinators.

In the next section we shall prove that claim 3 is indeed a general result.

Before we leave the three-agent example, we should discuss the fact that in a reasonable class of cases it appears to be better to make a coordinator junior; that is, hierarchy *ssc* is optimal. There may be interesting situations in which this is what we see in practice. Think of two firms engaged in a production or marketing alliance. Agent 3 is appointed to investigate synergies between the two firms. If alliance activities are regarded as less important than the specialist activities of the two firms, then 3 is made junior to the bosses of the two firms, agents 1 and 2.

Notwithstanding these kinds of situations, it is counterintuitive that a strategic planner should be low down in the hierarchy. Thus we are interested in finding conditions that knock out hierarchy *ssc*. Of course, if we are prepared to assume that strategic planners are highly productive ($v_c > 2v_s$), then we know from (4) that hierarchy *ssc* is dominated by hierarchy *css*, irrespective of probabilities. But the assumption that $v_c > 2v_s$ is quite strong.

If, instead, we are prepared to assume that coordinators have a lower probability of having an idea, then we know from result 1 that hierarchy

¹² Let $\Delta = V(css) - V(ssc)$, $\Delta' = V(sss) - V(ssc)$, and $\Delta'' = V(css) - V(ccc)$. If $p_c < p_s$, then $(\Delta/p_c) + [2\Delta'/(1-p_s)] > 0$, and hence one of Δ and Δ' must be strictly positive. If $p_s < p_c$, then $(\Delta/p_s) + [\Delta''/(1-p_c)] < 0$, and hence one of Δ and Δ'' must be strictly negative.

¹³ This intuition is only rough: the argument is more delicate in the middle ground, where v_c lies between $2v_s$ and v_s .

ssc is not optimal. In our opinion, this should be seen as the leading case. We think that the monotonicity condition $p_c < p_s$ is intuitively reasonable given our interpretation that the coordinator is a strategic planner. We have in mind that a strategic planner contemplates future directions for the organization. For example, a new direction might be to change the product line or location or to put more emphasis on foreign sales rather than on domestic sales. If such a new direction is found and implemented, then this will have major repercussions for all the other people working on these assets (formally, it will preclude the implementation of their ideas). Finding such a new direction is not a routine matter, however, and this is why the strategic planner's probability of having an idea may be low.

V. General Results

In this section we establish some general results about optimal organization forms. As promised, here is the generalization of claim 3 that will be the key to all our other results.

LEMMA. Consider an optimal hierarchy. If $p(A_j) < p(A_i)$, then for all assets $a_k \in A_i \cap A_j$, j appears above i on list L_k .

Proof. See the Appendix.

It may help to rehearse why this result is not as odd as it looks! It seems to suggest that the choice of hierarchy is determined solely by the probability of success and not at all by the value of an idea. But the lemma says nothing about which tasks should be undertaken. If a task has a low probability of success *and* a low value, no one should be doing it. In an optimal hierarchy, an agent whose task has a low probability of success, who according to the lemma should be senior, must be able to generate a commensurately high value to warrant his doing that task in the first place.

This observation about agents having the wrong task gives the clue to how the lemma is proved. Take an asset a_k , and suppose that agent i is senior to agent j on list L_k , even though $p(A_j) < p(A_i)$. We show that total expected value can be increased by making one of two changes to the organizational form. Either i is relatively unproductive ($v(A_i)$ is relatively low), in which case expected total value can be increased by switching i to task $t(A_j)$ and placing him just under j in seniority on all assets in A_j . Or j is relatively unproductive ($v(A_j)$ is relatively low), in which case expected total value can be increased by switching j to task $t(A_i)$ and placing her just under i in seniority on all assets in A_i .¹⁴

¹⁴ The merit of these two kinds of maneuvers is that one can keep track of how expected total value changes. By contrast, if we consider a third kind of change to the organizational form, the ostensibly more straightforward maneuver of simply switching the seniorities of

Our next result is a generalization of claim 1.

PROPOSITION 1. In an optimal hierarchy, crisscross arrangements are never optimal. That is, if j appears above i on list L_k for asset a_k , then there does not exist k' such that i appears above j on list $L_{k'}$ for asset $a_{k'}$.¹⁵

Proof. See the Appendix.

Notice that when $p(A_i) \neq p(A_j)$, proposition 1 follows immediately from the lemma. The heart of the proof, therefore, lies in showing that crisscross arrangements are also not optimal when $p(A_i) = p(A_j)$.

Of course, the leading case in which $p(A_i) = p(A_j)$ occurs when $A_i = A_j$; proposition 1 then says that, given two people with the same task, one should be senior to the other on all the assets they work on.

Our next result, proposition 2, deals with the more interesting case in which $A_i \subset A_j$, $A_i \neq A_j$; that is, j 's remit is broader than i 's. It generalizes result 1 and makes use of the following monotonicity condition. Call any task $t(A)$ for which $v(A) > 0$ a *productive* task. In an optimal hierarchy, everyone has a productive task.¹⁶

MONOTONICITY (M). For any two productive tasks $t(A)$ and $t(B)$, $A \neq B$, if $A \subset B$, then $p(A) > p(B)$.

As we argued at the end of Section IV, we believe that condition M is plausible: the probability of an idea is decreasing in an agent's span of control.

PROPOSITION 2. Assume condition M. Consider an optimal hierarchy. Suppose $A_i \subset A_j$, $A_i \neq A_j$. Then, for each asset $a_k \in A_i$, j appears above i on list L_k .

Proposition 2 follows directly from the lemma. It says that, under condition M, it is optimal for someone with a broad remit to be senior to someone with a narrow remit. However, the proposition covers only the case in which agents' remits can be nested. To this extent, it allows for organizational forms that may appear unusual. Consider the following example.

Example.—There are three agents and four assets. No single asset is productive (i.e., $v(\{a_i\}) = 0$ for all i). The only pairs of assets that are productive are $\{a_1, a_2\}$, $\{a_3, a_4\}$, and $\{a_2, a_3\}$. Assume $v(\{a_1, a_2\}) =$

i and j , then the people who lie between i and j (on list L_k) are affected in subtle ways and the impact on expected total value is too messy to be useful.

¹⁵ This can be generalized further. It is never optimal to have a "cycle of seniority" in which, e.g., 1 is somewhere senior to 2 (i.e., 1 appears above 2 on the list for some asset), 2 is somewhere senior to 3, ..., $i-1$ is somewhere senior to i , and i is somewhere senior to 1.

¹⁶ If someone had a task $t(A)$ for which $v(A) = 0$, he would not be able to generate any value. But such a hierarchy cannot be optimal, because instead he could be given a productive task $t(A')$ and, for each asset $a_k \in A'$, he could be placed at the bottom of list L_k so as not to block anyone else. He would then be able to generate value in the event that he has an idea but no one senior to him does.

TABLE 3

Asset a_1	Asset a_2	Asset a_3	Asset a_4
1	1 3	2 3	2

$v(\{a_3, a_4\}) = 16$, $p(\{a_1, a_2\}) = p(\{a_3, a_4\}) = \frac{1}{2}$ and $v(\{a_2, a_3\}) = 20$, $p(\{a_2, a_3\}) \approx 1$. (For all larger sets of assets, $v = 32$, say, but $p \approx 0$.)

Note that in this example the monotonicity condition M holds.

Now, it is straightforward to confirm that in this example the arrangement in table 3 is an optimal hierarchy: agent 1 is in charge of assets a_1 and a_2 , agent 2 is in charge of assets a_3 and a_4 , and agent 3 is a subordinate on assets a_2 and a_3 .¹⁷ What is somewhat unusual about this hierarchy is that agent 3 has two bosses: agent 1 on asset a_2 and agent 2 on asset a_3 . The arrangement in table 3 might be referred to as a nonpyramidal hierarchy (or, perhaps, as matrix management).

We can rule out nonpyramidal hierarchies if we are prepared to make a further assumption.

NESTEDNESS (N). Given two productive tasks $t(A)$ and $t(B)$ (i.e., $v(A)$ and $v(B)$ are both strictly positive), either $A \subset B$, $B \subset A$, or $A \cap B = \emptyset$.

In other words, condition N says that if a task $t(A)$ is productive, then any synergy involving one of the assets in A and a new asset requires the presence of the other assets in A too. The example does not satisfy condition N: take $A = \{a_1, a_2\}$ and $B = \{a_2, a_3\}$.

Condition N is strong. Note, however, that it is trivially satisfied in the two-asset case.

PROPOSITION 3. Assume conditions M and N. Consider an optimal hierarchy. If j appears above i on list L_k for some asset a_k , then j appears above i on every list on which i appears. That is, the hierarchy is a pyramid. Each agent has the same ranking on every asset he works on.

The proof of proposition 3 is direct. In an optimal hierarchy, i and j must both have productive tasks. It follows from condition N that $A_i \subset A_j$, $A_j \subset A_i$, or $A_i \cap A_j = \emptyset$. The last is impossible since j appears above i on some asset. If $A_i \subset A_j$, $A_i \neq A_j$, the conclusion of proposition 3 follows from proposition 2. If $A_i = A_j$, the conclusion follows from

¹⁷ To see why this hierarchy is optimal, note that the date 0 expected total value equals 21: agents 1 and 2 on average contribute 8 each, and agent 3's expected contribution is 5, given that he has access to a_2 and a_3 only if neither 1 nor 2 has an idea (which happens with probability one-fourth). If 3 were placed junior to 1 on assets a_1 and a_2 , his expected contribution would be only 4 (he would generate 16 when 1 did not have an idea and he did, which would happen with probability one-fourth). If 3 were placed senior on assets a_2 and a_3 , the total value would be just 20.

proposition 1. Finally, $A_j \subset A_i$, $A_j \neq A_i$ is inconsistent with proposition 2 since we know that i does not appear above j on list L_k .

Finally, we present a sufficient condition for a hierarchy to achieve first-best. The condition is a natural generalization of the condition $v_c > 2v_s$ that we discussed in Section IV. We say that $v(\cdot)$ is *conditionally superadditive* if, for any productive task $t(A)$ and any partition $A^1 \cup A^2 \cup \dots \cup A^p$ of A , $v(A) \geq v(A^1) + v(A^2) + \dots + v(A^p)$. (This is only “conditionally” superadditive because it applies just to subsets A for which $v(A) > 0$, i.e., for which the task $t(A)$ is productive.)

PROPOSITION 4. Assume condition N. If $v(\cdot)$ is conditionally superadditive, an optimal hierarchy achieves first-best.

Proof. See the Appendix.

Note that proposition 4 does not require the monotonicity condition M. But the nestedness condition N *is* needed. Consider again the example given prior to table 3, where $v(\cdot)$ is conditionally superadditive but condition N does not hold (though condition M does). The first-best is not achieved because in the event that, say, both 1 and 3 have ideas but 2 does not, it would be better if 3 were senior to 1 on asset a_2 given that 3 could generate 20 whereas 1 can generate only 16.

VI. Centralization versus Decentralization

One of the benefits of our model is that it provides a fairly clear-cut way of thinking about the terms centralization and decentralization. We say that an organization is more decentralized if it is more likely that a decision will be made by someone on the ground, that is, a local expert or specialist.

Return to the example from Section IV, of two symmetric assets, a_1 and a_2 , where there are just two tasks, specialization and coordination. We shall use this example to explore the trade-off between centralized, decentralized, and independent firms.

We can make our points most simply by considering the case of four agents, 1, 2, 3, and 4. (Everything we say generalizes to $n > 4$.) As in Section IV, let p_s denote a specialist’s probability of having an idea and v_s its value, and let p_c denote a coordinator’s probability of having an idea and v_c its value. Let us further assume monotonicity condition M: $p_c < p_s$. Then we know from propositions 1 and 2 that coordinators will be senior to specialists and crisscross arrangements are not optimal. One can further show that an optimal hierarchy is symmetric: there will be the same number of specialists working on each asset.¹⁸ In effect,

¹⁸ In Hart and Moore (1999), we show that this symmetry result is true for any value of n : in an optimal hierarchy, there is always an equal number of specialists working on each asset. (The only exception is the case in which n is odd and there are no coordinators, in which case the nearly symmetric hierarchy is optimal: one asset has an extra specialist.)

TABLE 4

HIERARCHY <i>C</i> (a Centralized Firm)		HIERARCHY <i>D</i> (a Decentralized Firm)		HIERARCHY <i>T</i> (Two Firms)	
Asset a_1	Asset a_2	Asset a_1	Asset a_2	Asset a_1	Asset a_2
1	1	1	1	1	2
2	2	2	2	3	4
3	3	3	4		
4	4				

then, there are only three candidates for an optimal hierarchy, illustrated in table 4.

The first two hierarchies, *C* and *D*, represent a single firm since both assets have the same ultimate boss, agent 1 (who can be interpreted as the owner of the assets, or possibly as the board of directors). Hierarchy *D* can be thought of as a more decentralized firm than hierarchy *C* because authority is more likely to be exercised by a specialist: in hierarchy *D*, 3 or 4 gets to exercise authority if 1 and 2 do not have ideas. The third possible hierarchy, *T*, represents two firms since assets a_1 and a_2 have different ultimate bosses: 1 is the boss of a_1 and 2 is the boss of a_2 .

The date 0 expected total values for the three hierarchies are, respectively,

$$V(C) = [1 - (1 - p_c)^4]v_c \quad (7)$$

$$V(D) = [1 - (1 - p_c)^2]v_c + (1 - p_c)^2 2p_s v_s, \quad (8)$$

and

$$V(T) = 2[1 - (1 - p_s)^2]v_s. \quad (9)$$

It is useful to think of the expressions in (7)–(9) as functions of v_c : $V(C)$ is proportional to v_c (there are no specialists in hierarchy *C*, so v_s does not appear), $V(D)$ is affine in v_c with a positive intercept but a smaller slope than $V(C)$, and $V(T)$ is independent of v_c (since there are no coordinators in hierarchy *T*). Hence, plotting $V(C)$, $V(D)$, and $V(T)$ against v_c , keeping v_s constant, we can get just two possible diagrams, depending on whether the $V(D)$ line lies above or below the intersection of the $V(C)$ and $V(T)$ lines (see fig. 1).

There is a critical value of p_c at which all three lines intersect:

$$p_c^* = 1 - \sqrt{1 - p_s}. \quad (10)$$

The term p_c^* is a function of p_s and lies strictly between zero and p_s .

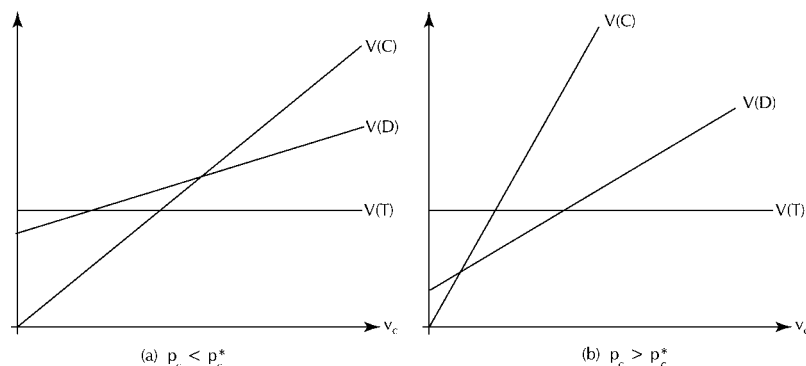


FIG. 1

Figure 1a pertains to $p_c < p_c^*$, and figure 1b pertains to $p_c > p_c^*$ (our monotonicity assumption $p_c < p_s$ is always maintained).

Organizational form is chosen to maximize V , and so only the upper envelope of the three lines is relevant. If $p_c < p_c^*$, it follows from figure 1a that, for low values of v_c (relative to v_s), hierarchy T is best; for moderate values of v_c , hierarchy D is best; and for high values of v_c , hierarchy C is best. In contrast, if $p_c > p_c^*$, it follows from figure 1b that hierarchy D is never optimal: for low values of v_c , hierarchy T is best; and for high values of v_c , hierarchy C is best.¹⁹

The conclusion is that, at least when p_c is not too close to p_s , a single centralized firm, a single decentralized firm, and two independent firms can all be optimal, depending on whether the value from coordination is high, moderate, or low.

It is worth remarking that although we think that our definition of decentralization—that decisions are more likely to be made by a specialist—is reasonable, it is not standard.²⁰ It is not uncommon for economists to define decentralization in terms of span of control. For example, according to Wernerfelt (1992), an organization in which many workers report to one boss is centralized (this is a flat hierarchy), whereas an organization in which there is a level of middle managers between the workers and the boss is decentralized (this is a steep hi-

¹⁹ In Hart and Moore (1999), we show that this nonconvexity in the choice of hierarchy is general for all n . Specifically, we show that if $p_c > p_c^*$, where p_c^* is given in (10), then a single decentralized firm is never optimal; i.e., there is never an interior solution for the optimal number of coordinators and specialists. A single decentralized firm either is dominated by a fully centralized firm with no specialists (if v_c is high enough: $v_c \geq 2v_s$ is a sufficient condition) or is dominated by two firms with no coordinators. However, if $p_c < p_c^*$, a single decentralized firm is optimal when v_c lies in the intermediate range between $2v_s$ and $2v_s p_c^* / p_c$.

²⁰ A related definition may be found in Aghion and Tirole (1997).

erarchy). We would say, however, that the presence of the middle managers makes the organization *more* centralized, not less, because the probability that a worker (specialist) makes a decision is reduced.

Note that we are not saying that one definition is right and the other wrong.²¹ However, we do feel that it may be useful to have a notion of decentralization based on the likelihood that local experts make a decision.

VII. Interpretations and Foundational Issues

In this section we present some interpretations of the model. These interpretations provide contexts in which it may be reasonable to assume that decision making is carried out according to a hierarchical structure, the parties cannot renegotiate at date 1, and yet organizational form is chosen efficiently at date 0. That is, they provide a (limited) justification for focusing on hierarchies—and not considering more general mechanisms for decision making.

A. Imagine that the value $v(A)$ from task $t(A)$ represents a common benefit, earned by all individuals 1, ..., n ; that is, we are in a “teams” situation. However, time is very short at date 1 and individuals must act independently. (In contrast, there is plenty of time to set up the decision-making process at date 0.) For example, the individuals might represent the officers in an army that is engaged in a war. The assets correspond to the troops involved in the different battles and the decisions on how each battle is to be fought. The v 's represent the values from fighting battles in certain ways. Officers may have different ideas about what the soldiers should do. (Given army discipline, the soldiers will do what they are told. However, only one decision or plan can be implemented for each battle.) Suppose that time is so short that not only can the officers not get together to discuss their various ideas (they are in the heat of battle), but each officer does not even know whether other officers have any ideas; he knows only his own idea. Then in equilibrium each officer with an idea will issue an instruction to the soldiers. Clearly the soldiers require some way to decide between conflicting instructions. It is plausible that the mechanism used will be a hierarchy: the troops are told to obey the general if he issues an instruction; the colonel if he issues an instruction, but the general does not; and so on.

Assumptions 1–8 are by and large reasonable in this context, and so our analysis provides a characterization of the optimal organizational form. Our analysis can explain why the most senior officers (generals)

²¹ There are also other reasonable definitions of decentralization. For example, one might say that an organization is decentralized if the authority to make certain decisions is held by lower levels of the organization, e.g., specialists. On this, see Bolton and Dewatripont (1995) and Hart and Holmstrom (2002).

are given the task of overseeing the whole war, whereas a colonel or a major might be responsible for a particular battle.

B. The “teams” interpretation of the model relies on the assumption that individuals do not know who else has had an idea (or the value of this idea); otherwise seniority does not matter since whoever is in authority will defer to the person with the best idea.²² So consider the same setup as above, but suppose instead that it is common knowledge as to who has had an idea and the value of this idea. However, assume now that v is a *private* benefit; that is, v accrues entirely to the person having the idea (if the idea is implemented; if it is not implemented, there is no value). Then a senior person with an idea will never defer to a junior person, since the senior person does not benefit from it. (We must also assume now that an individual with an idea who has access to some but not all of the assets corresponding to his task still gets a small benefit.) Suppose that there still is not time at date 1 for renegotiation with side payments. Then our assumption that decisions are made hierarchically is plausible. Moreover, if the parties have plenty of time at date 0 and can make side payments, it is reasonable to suppose that they will choose organizational form *ex ante* to maximize expected value (the Coase theorem). Again the model applies.

C. A different interpretation of the model can be given, based on a managerial view of the firm. Assume that the firm is a public company, and $v(A)$ is profit accruing to dispersed shareholders. Suppose that individuals 1, ..., n represent managers, who are motivated by power: to be specific, assume that each manager gets an ego boost from having his idea implemented (if he has one) but does not care about money at all. Then managers will never defer to each other if they have an idea and no *ex post* renegotiation will occur (there is no medium of exchange for compensating managers).²³ Assume finally that the firm is set up by an entrepreneur at date 0 who wishes to sell off his entire equity stake—representing claims on future profit—for the highest price (he then retires). Then the entrepreneur will choose hierarchical structure and task assignment to maximize the expected value of the date 1 profit. (In this interpretation the highest person in the chain of com-

²² This statement ignores some potentially serious computational issues. Figuring out *ex post* how to allocate ideas to assets may be very difficult. In fact, Merlone (2003) shows that, under certain conditions, the problem of achieving the first-best in the model of this paper is “nondeterministic polynomial-time hard”; he interprets a hierarchy as a particular heuristic to solve this problem.

²³ In this interpretation of the model (and in the interpretation in paragraph B), one might ask whether it is reasonable to suppose that a manager will *ever* defer to a subordinate; after all, won’t he always have a poor idea that he would prefer to implement? Assume, however, that there is a cost of implementing an idea and that the manager’s ego boost (private benefit) varies with the value of the idea. Then, low-value ideas will be abandoned; i.e., the manager will defer if his idea has a low enough value.

mand on an asset represents the board of directors since there are no active owners.)

D. A view related to that in paragraph C can be given if we are prepared to adopt a non-Bayesian perspective. Assume that individuals 1, ..., n care only about money and $v(A)$ represents a common benefit. (So we are in a teams situation again.) However, suppose that each individual who has an idea stubbornly thinks that his idea is the best—and is impervious to argument. Then individuals will never defer to each other and no ex post renegotiation will occur. Assume finally that the individuals are “sensible” at date 0 and choose organizational form to maximize expected value, taking into account their future stubbornness. Then the model applies.

Although the above interpretations are all quite tentative or restrictive, they do provide some justification for the model, in the sense that they specify conditions under which more elaborate mechanisms than a hierarchy may be infeasible or very costly. Note that an entirely different kind of justification can be given: we have seen that in some cases a hierarchy can do remarkably well—in fact achieve first-best (see proposition 4). Under these conditions, of course, one does not have to worry about more elaborate mechanisms.

VIII. Further Discussion

We have established several results. First, we have shown that, given the assumption that the probability of an idea is decreasing in the set of assets being looked after, individuals with a broad remit, that is, whose tasks cover a large subset of assets, should appear higher in the chain of command than those with a narrow remit. In other words, big thinkers or coordinators should be senior to small thinkers or specialists. Second, we have shown that “crisscross” hierarchies are never optimal; that is, if individual i appears above individual j on one asset, j will not appear above i on another asset. Third, if synergies are nested, the optimal hierarchy is a pyramid. Fourth, large synergies across assets or decisions lead to large, centralized firms, whereas smaller synergies lead to decentralized or independent firms.

It would obviously be desirable to see whether these results are compatible with the internal organization of actual firms. This is a task for future research, but it is worth noting that our finding that strategic planners are placed at the top of an organization does seem descriptive of reality. Also, our model is consistent with the observation that firms are becoming more decentralized over time (see Rajan and Wulf 2003). It is plausible that the increases in flexibility of modern technology have reduced the importance of coordination or synergy; that is, in terms of the model of Section IV, v_c has fallen relative to v_s . According to our

model, this would be expected to lead to greater decentralization or to the creation of independent firms.

To obtain our results, we have had to make some quite strong assumptions. For example, we had to assume that a coordinator requires access to all the assets corresponding to his task in order to achieve value from an idea. We suspect that our results could be generalized to the case in which someone with a broad remit requires most, but not all, of the assets under his control to generate value. However, if coordinators *always* have ideas about small subsets (e.g., about individual assets), our results do change. In this case coordinators are not really coordinating; rather, they are multifaceted specialists, doing several tasks at once. For example, consider the two-asset case and suppose that there are two individuals, who think about both assets but have ideas about only one of them: specifically, individual 1 has an " a_1 " idea with probability one-half, in which case he generates 100 if he has access to a_1 ; and an " a_2 " idea with probability one-half, in which case he generates 80 if he has access to a_2 . The opposite is true for individual 2. Then a crisscross hierarchy achieves first-best: individual 1 should be senior on a_1 and junior on a_2 , and vice versa for individual 2. One interpretation of this is that each individual has a full-time job and a part-time job: he is senior in the first and junior in the second. Such arrangements are not uncommon, at least informally. For example, it is not unusual for someone who is the head of one division of a company to have to get approval from another division head if he has an idea that encroaches on the second division head's territory.

While we make no pretense that our results are entirely general, we do think that our model captures a significant aspect of organizations, namely, the fact that some people engage in strategic planning and others in more focused tasks, and it matters where each group is placed in the organization. Although strategic planners may only occasionally come up with major ideas, these ideas, when implemented, will necessarily have an impact on most, if not all, of the individuals working in the organization. In this respect, our assumption that a coordinator with an idea requires, and crowds out, access to all the assets he is working on may not be that unreasonable.

We turn now to some ways in which our work could be extended. First, it is desirable to endogenize the number of assets m and the number of individuals n ; this could be done by supposing that there is a competitive market and additional workers and assets can be purchased at market prices. It is also interesting to carry out various comparative statics exercises. In an earlier version of this paper we analyzed the two-asset case and considered what happens to the optimal hierarchy as the number of agents increases. We found that if coordination occurs at all, it increases (see Hart and Moore 1999, sec. 4, results 2, 3). How-

ever, there are other margins of interest. One could ask what happens as the number of assets or decisions increases along with the number of agents. Depending on the nature of synergies, expansion might then take the form of the hiring of further specialists rather than increased coordination.

Second, it would be useful to incorporate *ex ante* effort or investment choices. We have supposed that the probability of having an idea is exogenous and is independent of effort. In this respect, our model differs sharply from the “property rights” theory of the firm (for a summary, see Hart [1995]). In that literature, it is typically assumed that the parties can costlessly renegotiate *ex post*, and the role of organizational form is to provide parties with the right incentive to make *ex ante* investments. In this paper we have gone to the other extreme: we have ruled out *ex post* renegotiation and ignored *ex ante* incentives. In fact, we have taken the view that hierarchy is a substitute for renegotiation. The truth obviously lies somewhere in between.

We end by mentioning two potential applications of the model. The first is to understand how economies differ during periods of national emergency (“war”) as opposed to periods of normality (“peace”). There is a sizable literature arguing that centralization becomes more desirable in time of war, but it has been hard to formalize this idea.²⁴ Our model may be relevant here, given that it predicts that centralization is optimal when coordination benefits are sufficiently great.

Second, ever since Chandler’s (1962) work, economists have been interested in the choice between the U form and the M form.²⁵ To understand this choice in our model, consider a stylized automobile company. Suppose that there are four assets, a_1 , a_2 , a_3 , and a_4 , where a_1 represents an asset for producing luxury cars, a_2 represents an asset for producing family cars, a_3 represents an asset for marketing luxury cars, and a_4 represents an asset for marketing family cars. One way to structure the automobile company is to combine a_1 and a_2 in one division and a_3 and a_4 in another division. This corresponds to the U form (organization by function). A second way to structure the company is to combine a_1 and a_3 in one division and a_2 and a_4 in another. This corresponds to the M form (organization according to self-contained units). These two arrangements differ in task design, as well as in design of hierarchy. It would be interesting to use the model to analyze the trade-off between the two.

²⁴ For an interesting attempt in this direction, see Bolton and Farrell (1990).

²⁵ For a recent formalization, see Maskin, Qian, and Xu (2000). For a discussion of how the U form and M form compare in their ability to solve coordination problems, see Argyres (1995).

Appendix

Proofs of the Lemma and Propositions 1 and 4

The proofs of the lemma and of proposition 1 both make use of the following argument.

Consider an optimal organizational form and two individuals i and j such that $A_i \cap A_j$ is nonempty. To ease notation, let $p_i = p(A_i)$ and $p_j = p(A_j)$. By assumption, $0 < p_i < 1$ and $0 < p_j < 1$. Define X as the expected surplus conditional on neither i nor j having an idea, X_i as the expected surplus conditional on i having an idea but not j , X_j as the expected surplus conditional on j having an idea but not i , and X_{ij} as the expected surplus conditional on both i and j having an idea. That is, the (unconditional) expected surplus equals

$$V = (1 - p_i)(1 - p_j)X + p_i(1 - p_j)X_i + (1 - p_i)p_jX_j + p_ip_jX_{ij}. \quad (A1)$$

Since we are at an optimum, we know that no organizational form can generate surplus strictly in excess of V .

Now consider four possible changes to the organizational form. Each alternative entails changing the seniority/task of either i or j but not both people together (throughout, we leave everyone else's tasks and seniorities unchanged). These changes will allow us to establish some useful inequalities.

First, lower j 's seniority so that she is at the bottom of every list on which she appears. If we ignore j 's contribution, then expected surplus is given by $(1 - p_i)X + p_iX_i$. Hence, since V is the maximum surplus that can be obtained,

$$(1 - p_i)X + p_iX_i < V. \quad (A2)$$

The inequality is strict because, although j is now extremely junior, she sometimes contributes $v(A_j)$: when she has an idea and none of the people senior to her has an idea. (If her task were unproductive, i.e., if $v(A_j)$ were equal to zero, she could be allocated a productive task instead and still kept junior.)

Second, give j the same task as i and move her immediately below i in seniority on the lists pertaining to assets A_i , including those assets not in A_j (she no longer appears on the lists pertaining to those assets in A_j that are not in A_i). The new expected surplus cannot exceed the maximum:

$$(1 - p_i)^2X + [1 - (1 - p_i)^2]X_i \leq V. \quad (A3)$$

Third and fourth, the changes to the organizational form that led to (A2) and (A3) can be repeated reversing the roles of i and j :

$$(1 - p_j)X + p_jX_j < V \quad (A4)$$

and

$$(1 - p_j)^2X + [1 - (1 - p_j)^2]X_j \leq V. \quad (A5)$$

Inequalities (A2)–(A5) will be of use in proving both the lemma and proposition 1.

Proof of the Lemma

Suppose that the lemma is false. That is, suppose that $p_j < p_i$ but there exists some $a_k \in A_i \cap A_j$ such that i appears above j on list L_k . Then, conditional on i having an idea, j cannot generate any value because i is senior to j on one of her assets, a_k . That is, i blocks j . Worse, when j also has an idea, not only is she

blocked, she in turn may block someone else junior to her from generating any value. In short, conditional on i having an idea, j 's presence in the organization serves only to reduce expected surplus:

$$X_{ij} \leq X_i. \quad (\text{A6})$$

We now show that (A6) is inconsistent with (A1)–(A5), given that $p_j < p_i$.

From (A1) and (A6),

$$V \leq (1 - p_i)(1 - p_j)X + p_j(1 - p_i)X_j + p_i X_i. \quad (\text{A7})$$

Combining (A2) and (A7), rearranging, and dividing by $p_j(1 - p_i)$, we obtain

$$X < X_j. \quad (\text{A8})$$

Combining (A4) and (A7), rearranging, and dividing by p_i , we obtain

$$p_j(X_j - X) < X_i - X,$$

which from (A8) implies

$$X < X_i. \quad (\text{A9})$$

Taking a convex combination of $(1 - p_j + p_i)/(2 - p_j)$ times (A3) plus $(1 - p_i)/(2 - p_j)$ times (A5), we obtain

$$(1 - p_i)(1 - p_j)X + p_j(1 - p_i)X_j + p_i X_i + \frac{p_i(1 - p_i)(p_i - p_j)}{2 - p_j}(X_i - X) \leq V. \quad (\text{A10})$$

Combining (A7) and (A10) and dividing by $p_i(1 - p_i)/(2 - p_j)$, we obtain

$$(p_i - p_j)X_i \leq (p_i - p_j)X,$$

which, since $p_j < p_i$, contradicts (A9). QED

Proof of Proposition 1

As noted in the text, proposition 1 follows immediately from the lemma when $p_i \neq p_j$. Hence we need consider only the case $p_i = p_j = p$, say.

Suppose that the proposition is false: in an optimal hierarchy, j appears above i on some list and i appears above j on some other list. Then when both agents have an idea, they block each other: neither can generate any value. Worse, they may block someone else junior to them. Hence X_{ij} is bounded above by X_p , X_j , and X . Without loss of generality, we may assume that $X_j \leq X_i$. So

$$X_{ij} \leq X_j \leq X_i \quad (\text{A11})$$

and

$$X_{ij} \leq X. \quad (\text{A12})$$

Combining (A1) and (A3) with $p_i = p_j = p$, rearranging, and dividing by p , we obtain

$$(1 - p)(X_i - X_j) + p(X_i - X_{ij}) \leq 0.$$

This is consistent with (A11) and (A12) only if

$$X_{ij} = X_j = X_i \leq X. \quad (\text{A13})$$

Given the equalities in (A13), we conclude from (A1), with $p_i = p_j = p$, that

$$(1 - p)^2 X + [1 - (1 - p)^2] X_i = V.$$

But this is consistent with (A2) and $p_i = p$ only if $X < X_i$, which contradicts the inequality in (A13). QED

Proof of Proposition 4

We shall in fact prove a slightly stronger result than proposition 4. We shall show that, given any allocation of productive tasks across the n agents (i.e., tasks that each generate a strictly positive value), a pyramidal hierarchy maximizes value in all ex post events. The ex ante first-best is thus achieved via some pyramidal hierarchy through the optimal allocation of tasks.

The proof is by induction on n . The proposition holds trivially for $n = 1$.

Consider $n \geq 2$ agents and any allocation of productive tasks. Select one agent—let us call her agent n —whose task $t(A)$ is such that there is no one with a broader remit (i.e., there is no one whose task $t(\hat{A})$ is such that \hat{A} strictly includes A). By our nestedness assumption N, the other $n - 1$ agents can be partitioned into two sets, J and J' : in J , everyone has a task $t(\hat{A})$ such that $\hat{A} \subseteq A$; in J' , everyone has a task $t(\hat{A})$ such that $\hat{A} \cap A = \emptyset$.

From the induction hypothesis, in the absence of agent n , a pyramidal hierarchy, H^{n-1} , say, comprising the other $n - 1$ agents would maximize value in all ex post events. Construct a larger pyramidal hierarchy, H^n , comprising all n agents, which is the same as H^{n-1} except that agent n is placed in seniority above everyone else on each of the assets in A .

Consider the event in which agent n has an idea: event E , say. In this event, by the conditional superadditivity of $v(\cdot)$, value is maximized if agent n implements her idea, irrespective of which agents in J (or J') have ideas (agent n 's implementation of her idea does not interfere with the implementation of ideas by agents in J'). But given that agent n is senior on all the assets in A , hierarchy H^n thus always maximizes value conditional on E .

Now consider the complementary event in which agent n does not have an idea: event E' . Conditioning on E' has no effect on the values of any of the other $n - 1$ agents' ideas. Hence, conditional on E' , hierarchy H^n always maximizes value (given that the subhierarchy H^{n-1} always maximizes value in the absence of agent n).

In sum, hierarchy H^n maximizes value in all events. QED

References

- Aghion, Philippe, and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *J.P.E.* 105 (February): 1–29.
- Argyres, Nicholas S. 1995. "Technology Strategy, Governance Structure and Interdivisional Coordination." *J. Econ. Behavior and Organization* 28 (December): 337–58.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 1999. "Informal Authority in Organizations." *J. Law, Econ., and Organization* 15 (April): 56–73.
- Bolton, Patrick, and Mathias Dewatripont. 1994. "The Firm as a Communication Network." *Q.J.E.* 109 (November): 809–39.
- . 1995. "The Time and Budget Constraints of the Firm." *European Econ. Rev.* 39 (April): 691–99.

- Bolton, Patrick, and Joseph Farrell. 1990. "Decentralization, Duplication, and Delay." *J.P.E.* 98 (August): 803–26.
- Calvo, Guillermo A., and Stanislaw Wellisz. 1979. "Hierarchy, Ability, and Income Distribution." *J.P.E.* 87, no. 5, pt. 1 (October): 991–1010.
- Chandler, Alfred D., Jr. 1962. *Strategy and Structure: Chapters in the History of the Industrial Enterprise*. Cambridge, MA: MIT Press.
- Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production." *J.P.E.* 108 (October): 874–904.
- Harris, Milton, and Artur Raviv. 2002. "Organization Design." *Management Sci.* 48 (July): 852–65.
- Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Oxford Univ. Press.
- Hart, Oliver, and Bengt Holmstrom. 2002. "Vision and Firm Scope." Manuscript, Harvard Univ.
- Hart, Oliver, and John Moore. 1999. "On the Design of Hierarchies: Coordination versus Specialization." Working Paper no. 7388 (October), NBER, Cambridge, MA.
- Keren, Michael, and David Levhari. 1979. "The Optimum Span of Control in a Pure Hierarchy." *Management Sci.* 25 (November): 1162–72.
- Lando, Henrik. 2004. "On the Choice between Such Simple Mechanisms as Individual Ownership and the First-Come First-Serve Rule When Renegotiation Is Costly." Manuscript, Copenhagen Bus. School.
- Maskin, Eric, Yingyi Qian, and Chenggang Xu. 2000. "Incentives, Information, and Organizational Form." *Rev. Econ. Studies* 67 (April): 359–78.
- Merlone, Ugo. 2003. "Organizations as Allocators of Human and Material Resources." Manuscript, Turin Univ.
- Radner, Roy. 1992. "Hierarchy: The Economics of Management." *J. Econ. Literature* 30 (September): 1382–1415.
- Rajan, Raghuram G., and Julie Wulf. 2003. "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies." Manuscript, Grad. School Bus., Univ. Chicago.
- Rajan, Raghuram G., and Luigi Zingales. 2001. "The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms." *Q.J.E.* 116 (August): 805–51.
- Roberts, Kevin, and Martin L. Weitzman. 1981. "Funding Criteria for Research, Development, and Exploration Projects." *Econometrica* 49 (September): 1261–88.
- Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings." *Bell J. Econ.* 13 (Autumn): 311–23.
- Sah, Raaj K., and Joseph E. Stiglitz. 1986. "The Architecture of Economic Systems: Hierarchies and Polyarchies." *A.E.R.* 76 (September): 716–27.
- Tanenbaum, Andrew S. 2001. *Modern Operating Systems*. 2nd ed. Upper Saddle River, NJ: Prentice Hall.
- Van Zandt, Timothy. 1999. "Real-Time Decentralized Information Processing as a Model of Organizations with Boundedly Rational Agents." *Rev. Econ. Studies* 66 (July): 633–58.
- Wernerfelt, Berger. 1992. "The Structure and Scope of Firms and Economies of Scale in Contracting." Manuscript, Sloan School Management, Massachusetts Inst. Tech.
- Williamson, Oliver E. 1967. "Hierarchical Control and Optimum Firm Size." *J.P.E.* 75 (April): 123–38.